Summary of Public Feedback Received on the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) Proof of Concept

(2023-05-09)

Summary

The Census Bureau published a Proof of Concept for the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) on January 31, 2023. The Detailed DHC-A will provide population counts and sex and age statistics for detailed racial and ethnic groups and American Indian and Alaska Native tribes and villages. The Proof of Concept outlined how the product's differentially private algorithm, called SafeTab-P, uses an adaptive design that determines the amount of data racial and ethnic groups receive based on group size and geography level while ensuring sufficient confidentiality protections. The 30-day public feedback period on the Proof of Concept concluded on March 2, 2023.

The Census Bureau received 16 comments during the public comment period and two after the public comment period closed. Of these 18 comments, three were out of scope for this analysis, as their content was unrelated to the Proof of Concept (e.g. providing feedback on geographic boundaries, updates needed in future censuses to accurately capture hard-to-count communities, and the minimum categories for race and ethnicity collection). The remaining 15 comments were included in this analysis, with 11 requesting expanded guidance on how to use the data provided in the Detailed DHC-A.

Many commenters appreciated the Census Bureau prioritizing the publication of the population counts for an expanded number of detailed groups and acknowledged the benefits of the adaptive design to provide detailed data while protecting respondents' confidentiality. Additionally, commenters were pleased with the inclusion of tract and place level data, as well as population thresholds below 100 for publication. However, there were numerous concerns over how to aggregate data, such as collapsing sex by age data to make geographies comparable or to create custom geographies. Commenters asked for guidance on how to determine margins of error for custom aggregations as well as requesting guidance for the creation of derived statistics, such as percentages and ratios, and their respective quality indicators.

All comments were provided through the 2020DAS@census.gov email address. Additional feedback on the Detailed DHC-A was provided through engagement with stakeholders, Census advisory groups, and through formal tribal consultations, which is not contained in this summary. The written comments are summarized here and the full text of in-scope comments are included in Appendix A.

Background

This document provides a summary of the comments received for the Detailed DHC-A Proof of Concept. The Proof of Concept was released on January 31, 2023, and the deadline for public feedback was March 2, 2023. The two comments received after March 2, 2023 are also included in the summary provided.

Comments by Type of Affiliation

The comments received during this period represent a cross-section of data users, including federal, state, and local governments, academic institutions, and non-profits organizations (Table 1). Non-profit

organizations provided the most comments (6 comments), followed by local governments (4 comments), academic institutions (3 comments), and state and federal governments (1 comment each).

Table 1. Comments by Type of Affiliation

Type of Affiliation	**Count**
Federal government	1
State government	1
Local government	4
Academic institution	3
Non-profit	6

Geography

Two commenters expressly stated that they were pleased with the addition of tract level data as the lowest level of geography available, matching the 2010 Census data release. Eight commenters expressed having use cases for tract level data.

Three comments referenced geographies not available in the Detailed DHC-A. All comments indicated that tract level data were not always granular enough for their use cases. Two of these comments referred to needing block group data to identify and target resources to communities and areas of concern, while the other referenced needing block and block group data to estimate populations for areas of study.

Commenters' Concerns

The 15 comments received covered various concerns that were not mutually exclusive. Aggregating noisy counts from the Detailed DHC-A (12 comments) and the minimal guidance provided about using the Detailed DHC-A data (11 comments) were the most common concerns reported. Six commenters were concerned about regional group classifications and how these were determined. Another six commenters were concerned about the various types of thresholds applied in this product, particularly at sub-state geographies. Five commenters expressed concern about the visibility of margins of error and two were concerned about the utility of post-differential privacy suppressions. Lastly, the metrics, privacy-loss budget, the Detailed DHC-B, and the decrease in tables compared to 2010 data releases were all mentioned once, respectively. See Table 2 for an overview of commenter concerns.

Table 2. Commenter Concerns

Concern	Count
Aggregation	12
Guidance	11
Regional group classification	6
Thresholds	6
Visibility of margins of error	5
Suppression	2
Metrics	1
Privacy-loss budget	1
Detailed DHC-B	1
Fewer tables than in 2010	1

Aggregation

Aggregation was the most common concern, with 12 commenters stating that the current guidance, quoted below, was insufficient.

> *…. we caution data users against aggregating data. This includes creating new regional groups from detailed groups, combining lower-levels of geography to create higher-levels of geography, combining sex by age data, etc. We strongly recommend using the published noise infused counts associated with your geographic level of interest when available.*

Commenters expressed that they must aggregate data for their work (such as creating neighborhood data from tract level data), that larger margins of error are acceptable as long as they know what the margins of error are, and that data users will aggregate even if cautioned against it. Of the 12 comments that touched on aggregation, 10 stated that the Census Bureau needs to offer guidance on aggregating because data users will do so once the data are available, regardless of the implications for data accuracy.

Guidance

Table 3. Guidance Concerns

Concern	Sub Concern	Count
Guidance		11
	Aggregation	10
	How-to (general)	8
	Costs/benefits (general)	5
	Geographies	6
	Sex by age	5
	Derived measures (percentages, ratios, etc.)	6
	Comparisons	5
	Previous censuses	5
	American Community Survey (ACS)	5
	Suppression	6
	Alone larger than alone or in any combination	6
	Negative counts	2
	Communicating to lay audience	2

As previously mentioned, 11 commenters were concerned that insufficient or only technical guidance would be offered for using the Detailed DHC-A data. Ten of the 11 comments about guidance expressed that current guidance on data aggregation for the Detailed DHC-A was insufficient. Eight commenters asked for detailed and simple guidance on how to use aggregated data from the Detailed DHC-A. This included: 1) how to create aggregate measures with the least noise possible; 2) how to measure aggregated error; and 3) how to understand and communicate those aggregated outcomes – all for lay audiences who may not fully understand the use of margins of error in a decennial product. Additionally, five commenters asked for general guidance on the costs and benefits of aggregating. These commenters stated that it was important to ensure that data users understand fully what custom aggregations mean for data accuracy and when data users may want to avoid using custom aggregations. Six commenters mentioned wanting guidance specifically for aggregating geographies, and 5 mentioned guidance specifically for aggregating sex by age data.

Six of the 11 commenters requested further guidance on derived measures, such as percentages and ratios. Specifically, there were requests for guidance on the creation of quality indicators for derived measures, such as coefficients of variation and margins of error. These concerns overlapped with requests for guidance on how to use these indicators for comparison purposes. There was concern that data users will create derived statistics and compare to other data sources, such as the American Community Survey (ACS), without understanding the comparability of statistics gathered from various sources. Commenters recommend the Census Bureau provide tools to enable data users to create and understand quality indicators for derived measures.

Along the same lines, commenters requested guidance on how to compare Detailed DHC-A data to previous censuses and American Community Survey data. They indicated that previous guidance to "compare with caution" was insufficient and that guidance will be needed for comparing detailed

groups to other sources. Specifically, guidance will be needed for comparing data with margins of error created with different confidence intervals (95% for the Detailed DHC-A versus 90% for the American Community Survey), as well how the detailed groups compare to previous releases. Commenters specifically requested the Census Bureau provide tools to help them create comparison indicators and test significance between data, as is available for the American Community Survey.

Guidance on suppression and suppressed values came up six times. Two commenters requested that suppression not be applied and that guidance be offered instead. Other commenters requested easily accessible guidance informing data users that a suppressed alone value may have a non-suppressed paired alone or in any combination count and that data users should be encouraged to use this count instead if available. One commenter requested guidance be offered for negative counts they may encounter through "backing" into them when there are enough non-suppressed data cells to determine the suppressed count was negative in value.

Lastly, two commenters requested guidance on how to communicate the uses, challenges, and comparisons of the Detailed DHC-A data with lay audiences, to ensure that community organizations understand the data. Specifically, they requested language that is understandable and outside of technical documentation for topics such as the adaptive design determining how much data are produced for groups.

Regional Group Classification

Six commenters expressed concern about the Census Bureau's regional aggregations, specifically how various Asian detailed groups are classified. The following detailed groups were brought up as potentially misclassified:

- Hmong is classified as East Asian but should be classified as Southeast Asian (6)
- Indo-Chinese, Malay, and Timorese are classified as Other Asian but should be classified as Southeast Asian (4)
- Urdu is classified as Other Asian but should be classified as South Asian (4)
- Guyanese Indian is classified as Some Other Race but may be referencing people of Asian Indian descent in Guyana such as the Indo-Guyanese (4)
- Multiracial codes such as Amerasian and Eurasian should be tabulated as part of the Asian in combination population (4)

All six commenters raising concerns about misclassifications encouraged the Census Bureau to continue to engage with stakeholders of various backgrounds throughout all stages of development of racial and ethnic coding and classification in future endeavors.

Thresholds

Table 4. Threshold Concerns

Concern	Sub Concern	Count
Thresholds		6
	Sub-state minimum population counts	4
	Sex by age data	1
	True zeros	1

As previously mentioned, six commenters expressed concerns about the thresholds for receiving data. Four expressed concerns about the minimum population counts used at sub-state levels for detailed and regional groups. These commenters were concerned that the thresholds of 22 for detailed groups and 94 for regional groups were too high, which would result in many sub-state geographies having no data for detailed groups *or* their respective regional group within the same geography, potentially leaving data users without needed data. However, one commenter expressed that the thresholds of 22 and 94 were an improvement from 2010 when Asian, Native Hawaiian and Other Pacific Islander, and Hispanic origin groups had a threshold of 100.

One commenter was concerned that the threshold for 23 age category sex by age tables at sub-state geography levels is too high at 20,000 for detailed groups.

Another commenter expressed concern that too many true zeros will be published, affecting the integrity and usefulness of the data.

Visibility of Margins of Error

Five commenters expressed concern that the margins of error for counts will not be visible when looking up data and will only be available in technical documentation, meaning many data users will not know what they are.

Suppression

Two commenters were concerned that suppression done for non-confidentiality protection reasons, meaning the suppression of negative numbers and alone counts greater than their paired alone or in any combination counts, is a net-negative for the Detailed DHC-A data. Specifically, they were concerned that suppression makes the data more difficult for data users to understand and use. Both commenters preferred the Census Bureau recommends guidance for how to understand and interpret demographically unreasonable counts rather than use suppression. One commenter indicated that because this suppression does not happen for confidentiality purposes that a Freedom of Information Act request could potentially make these data available, regardless.

Metrics

One commenter expressed concern that the metrics released with the Proof of Concept were insufficient to determine the usability of the data. Specifically, the commenter indicated that knowing the number of geographies that have one or more cells suppressed and having an indicator of how often race or ethnicity groups have a summed state total at least 25% larger than their national total would help gauge whether the bias created by suppressing negative numbers is a problem for data quality.

Privacy-loss budget

One commenter expressed that the difference between the privacy-loss budget for the Proof of Concept for the Detailed DHC-A and the privacy-loss budget for the 2020 Census State Redistricting (Public Law 94-171) Summary File is noticeable and may require public justification.

Detailed DHC-B

One commenter expressed concern that the suppression of negative counts in the Detailed DHC-B will have a greater impact as households are smaller and more likely to be within a range where the margins of error can create negative counts if the Detailed DHC-B uses the same privacy-loss budget as the Detailed DHC-A.

Fewer tables than 2010

One commenter expressed concern that the Detailed DHC-A and Detailed DHC-B will provide fewer tables than were offered from the 2010 Census.

Recommendations

Table 5 summarizes commenter recommendations. Most of the comments recommend detailed and easy to understand guidance on the Detailed DHC-A, specifically its uses and limitations. Detailed and simple guidance on the pros and cons of data aggregation was the most recommended type of guidance. The requests for guidance are covered in more depth in the summary section on commenter's concerns about Guidance. Six commenters recommended that the Census Bureau adjust the regional groupings of the specific detailed race groups mentioned in the summary section on commenter's concerns about Regional Group Classification. Five commenters recommended that the margins of error, including aggregated margins of error, be available with the count in the data products so people will have them readily available and will not have to calculate aggregated margins of error themselves if using the provided counts. Four commenters recommended that the Census Bureau continues and expands its outreach to a variety of communities, especially considering the recommendation to make changes to the regional classifications of detailed race groups. Two commenters suggested the Census Bureau eliminate post-processing suppression all together. For negative counts, one commenter recommended setting a minimum population threshold of one at the nation and state levels to eliminate most negative counts, and both recommended letting negative counts in sex by age data be visible. Releasing new metrics to better assess data quality, publishing age tables without sex data, decreasing the threshold to 3,000 for a 23 age category table at sub-state geographies and using the total national population as thresholds for all sex by age data, increasing the current sub-state thresholds to further decrease the risk of enumerated (or true) zeros being published and including tabulations for household and group quarters population data were all recommended once, respectively.

Table 5. Commenter Recommendations

Recommendation	**Count**
Release detailed guidance on Detailed DHC-A limitations and uses	11
Adjust regional groupings	6
Make margins of error visible in data products	5
Continue/increase outreach	4
Eliminate postprocessing suppression	2
Release new metrics	1
Offer age tables (not by sex)	1
Change thresholds for sex by age data	1
Increase publication thresholds	1
Add household and group quarters data	1

Conclusion

Overall, the feedback received did not focus on the SafeTab-P algorithm, rather most commenters focused on the need to make the Detailed DHC-A as simple and clear for data users to understand and use as possible. Feedback along these lines ranged from recommendations to create easy to understand

guidance for aggregation to recommending that the margins of error be easy to find and aggregate for data users. Only two comments focused on the SafeTab-P algorithm itself, with one commenter expressing concern that the theoretical values they calculated differed from the values presented in the Proof of Concept and concern that the privacy-loss budget could potentially make vulnerable populations easier to target in certain geographic locations, and the other suggested lowering thresholds for 23 category sex by age data in sub-state areas.

Appendix A
Full Text Compilation of Feedback Received

Feedback on the Proof of Concept for the Detailed DHC A Product Released

1. Jan Vink, Cornell Program in Applied Demographics

Attached is my feedback on the Detailed DHC-A proof of concept released on January 31st

Jan Vink
Extension Associate
Cornell Program on Applied Demographics

Feedback on detailed DHC-A proof of concept

Jan Vink
Cornell Program on Applied Demographics
Email: jkv3@cornell.edu

Thanks for the opportunity to provide feedback on the Detailed DHC-A proof of concept. Involving stakeholders in the process of the DAS development is a critical component of the development cycle.

I start with my conclusions and recommendation and will expand on some of those points throughout the document.

Conclusions and Recommendations

- There are some potential usability issues that cannot be gauged with the published materials. The Census Bureau should consider publishing another proof-of-concept with metrics and examples that provide more insights in potential usability issues. The current proof of concept is focused too much on average accuracy metrics and examples.
- The rationale for the Privacy-Loss Budget choices are not clear, and I urge the Census Bureau to provide stakeholders such rationale.
- The metrics in the "Count Only" worksheet are different from the theoretical values, and I recommend that the Census Bureau checks those metrics to make sure that when all observations are included in the metrics, the numbers are close to the theoretical values and the current differences can be explained.
- In several of the examples the count of the number of people in the oldest age group was smaller than the MOE. If you calculate CVs for the oldest age groups, you will probably end up with values that indicate great risk for use. The data for the oldest categories could especially be important for some use cases that want to deliver services to specific detailed or regional race groups.
- The suppression rules are confusing and reduce the usability, so please reconsider them. I can imagine just keeping a rule where all geographies with a total count after noise infusion less than the threshold are suppressed. I would recommend a threshold of 1 instead of 0 for the nation and states, as a 0 count might have special meaning for many data users.
- The number of geographies that have zero count in the CEF but are published with a count because the added noise exceeds the threshold are likely impacting the usefulness of the Detailed DHC-A. A slight increase of the thresholds might prevent many of these cases.
- The only person variables used for each race group are age and sex. I think separate counts for household/group quarters populations will be greatly missed.
- Detailed race is now also collected in the ACS. It would be good to know whether there are plans to do something with that data in the future to understand whether the Detailed DHC-A is a once in a decade insight into detailed race or whether we will get annual survey based updates in the future. This would put different criteria on fitness for use of the detailed DHC-A.

Expected outcomes and metrics in Detailed Summary Metrics provided

My email correspondence with William Sexton at Tumult confirmed that there is a direct relation between the spread parameter of the Discrete Gaussian distribution (SIGMA^2) used to generate the noise and the MOE, namely that SIGMA^2 = (MOE/1.96)^2

Based on the different values for MOE, I was able to calculate expected values corresponding with the three different MOE settings.

	Expectations from Discrete Normal Distribution with spread parameter SIGMA^2				
MOE	SIGMA^2	Mean Absolute Error	Prob (x > MOE)	Threshold	Prob (x >= threshold | CEF = 0)
3	2.343	1.177	1.99%	1	1/2.7
11	31.497	4.466	4.02%	22	1/7,996
50	650.77	20.352	4.77%	1	1/2.03
				94	1/4,050

If I look at the Average difference and Percent of counts outside MOE on the "Age by Sex" worksheet most values correspond rather well with those expectations- although at the place and tract level, the Percent of counts outside MOE seem to be often a little lower than the expected 4.02%. This might have to do with the way the detailed metrics are calculated, which is not exactly the same and not completely intuitive.

If I look at the "Total Count Only" worksheet however, the Average Difference for Nation and State tends to be lower than the expected 1.177 and the Average Difference for Sub-State higher than the expected 4.466. The percentages of counts outside the MOE are much lower than expected for all levels of geography. These differences ask for further analysis, especially given that for sub-state geographies the average is higher than expected but the percent of counts outside MOE is smaller, when they should move in the same direction.

Missing metrics

- To assess the usability of this proof-of-concept it is important to understand how many of the geographies have one or more cells that are suppressed.
- There are many zeroes in the original data that are subject to noise and thus "at risk" of getting published counts of detailed race groups that are not present in that geography. The more this happens the less useful the dataset becomes.
- Bias: Because negative numbers are suppressed it is likely that the sum of the states is larger than the nation and, even with the thresholds in place It could be that sub-state geographies add up to larger values then the states. I suspect that this is most severe for smaller localized race groups and American Indian Tribes. Maybe a metric like the number of race groups where the sum of the states is more than 25% higher than the national count would help gauging whether this would be problem or not.

Theoretically, with an MOE of 3 (used for detailed race groups at the state level), 1 out of 2.7 states that had a zero count in the CEF will get a positive count after noise injection. With an MOE of 50 (used for regional groups at the state level), 1 out of 2.03 states will get a positive count after noise injection.

For sub-state geographies, roughly 1 in 8,000 geographies with CEF counts of zero will exceed the threshold of 22 for detailed groups and 1 in 4,000 geographies will exceed the threshold of 94 for

regional groups where the CEF count was zero. This might sound like low odds, but given that there are over 70,000 tracts and hundreds of detailed groups, many groups will get published counts in places where no one was counted.

Rationale for Privacy-Loss Budget

Overview of Privacy-Loss Budgets for 2020 data products and 2010 based demonstration datasets

Product	PLB (Rho)
Detailed DHC-A proof-of-concept	19.776
2020 PL 94-171 (person tables)	2.56
August 2022 Demonstration Product – person tables	3.65
August 2022 Demonstration Product – Units tables	6.14
Sub-State Census Process indicators	0.1

Given that privacy loss ~ exp(rho), the difference between the PLB for this proof-of-concept and that of the redistricting data is striking. What various actions by bad actors are considered in the different data products? This product would be the first Census long-form count for potentially vulnerable population groups like those of Israeli origin. Could the DHC-A be abused to target certain neighborhoods with people of Israeli origin? Should the Disclosure Avoidance System be preventing these kinds of abuse?

It would greatly add to the transparency of the DAS development if the Census Bureau would explain the interpretation of Title 13 and the rationales that lead to the choices of rho for each of the data products.

Suppression rules

The proof of concept contains two suppression rules: one for negative counts and one for some cells on race alone.

Especially the latter rule on race alone is likely to confuse people and not many use cases will need to use the race alone count and the race alone or in combination count in the same application. I think it would be easy to explain in a guidance document that these instances are possible.

Guidance on the presence of negative numbers might be a bit harder to develop but needs to be done, as some of the negative numbers can be derived by calculating residuals in the age tables for any suppressed values. I would prefer the transparency and therefore publishing negative numbers in the age/sex tables.

Negative total counts do not contribute to the usefulness of the DHC-A and those race group/geography combinations should just not be published instead of being published with an X.

A total count of zero has a special meaning, so I would opt to not publish zero counts and make publication of totals only at the nation/state level start with 1.

2. IPUMS

Hello,

I'm submitting feedback on the DDHC Proof of Concept from IPUMS staff members.

Yours,
Dave Van Riper

--
David Van Riper
Director of Spatial Analysis
Minnesota Population Center
Institute for Social Research and Data Innovation
University of Minnesota
(he/him/his)

Feedback on the DDHC-A Proof of Concept

David Van Riper, Tracy Kugler, Steven Ruggles, and Jonathan Schroeder
IPUMS
Institute for Social Research and Data Innovation
University of Minnesota
2023-03-02

Introduction

The US Census Bureau released a Proof of Concept for the Detailed Demographic and Housing Characteristics A product (DDHC-A) in January 2023. This Proof of Concept provided examples of the output data for detailed racial, ethnic, and American Indian and Alaska Native tribal and village groups along with a document describing the examples and the algorithm used to create them.

We appreciate the opportunity to provide feedback on the both the examples and the documentation. Our feedback focuses on the use of data suppression and messaging about data aggregation. We strongly discourage the Bureau from suppressing data in the DDHC-A (or future DDHC products) due to "demographic unreasonableness," and we suggest that messaging about data aggregation should be less cautionary and more explanatory.

Do not suppress "unreasonable" data

The Bureau's Proof of Concept document describes four circumstances that result in data suppression in the published product:

1. Noise-infused counts for a particular {geographic unit – race/ethnic/AIAN group} are negative
2. The "Race alone" count exceeds the "Race alone or in combination" count for a particular {geographic unit – race/ethnic/AIAN group}
3. Detailed groups with populations < 22 and regional groups with populations < 94 for a particular sub-state geographic unit or AIANHH area will not have data published for that geographic unit
4. Counts not included in the product

We are particularly concerned about circumstances (1) and (2) because they hide useful data from the public unnecessarily. The stated reason for suppressing these cases is that they are not "demographically reasonable," but there is no explanation for why it is important or beneficial to withhold demographically unreasonable data from users. We suppose the aim is mainly to prevent negative reactions to the data quality, but the only meaningful way to mitigate such quality issues is to prevent them in the first place. Given the nature of the noise-infusion algorithm, some "demographically unreasonable" results are to be expected. Suppression instead compounds one problem by adding another, obscuring the underlying issue and restricting users' options, with no gain in privacy protection. To the extent that the unreasonable counts are a necessary byproduct of maintaining adequate privacy

protections, then we urge the Bureau to publish all the problematic data and clearly explain how and why the problems exist.

The next sections describe the two problematic circumstances in more detail.

Negative counts

The algorithm (SafeTab-P) used to produce the DDHC proof of concept infuses random noise into counts derived from the confidential data. If a count is relatively close to zero, it is possible that the noise infusion process may yield a negative privacy-protected count. The Bureau proposes to suppress all such instances in the published data "because negative counts are not demographically reasonable" (Proof of Concept, p. 18).

SafeTab-P was designed to produce noise-infused counts that fall within a specified margin of error (MOE) around the confidential counts 95% of the time. The MOEs, particularly those for detailed groups, are relatively small (Table 1). Thus, when the algorithm produces a negative count, data users can be certain that confidential count was close to zero.

Table 1. Margins of error by geography for detailed groups.

Geography	MOE
Nation	3
State	3
County	11
Place	11
Tract	11
AIANHH	11

Negative counts are clearly unreasonable, and there may be settings where the best analytical strategy is to ignore them, but publishing negative values would not prevent users from doing so. Meanwhile, suppressing negative values does prevent users from making any direct use of the suppressed data. With unsuppressed negative values, users could easily and flexibly aggregate cells into larger categories or geographic areas to produce more reliable, useable data. They could choose to aggregate enough cells to eliminate negative values if that is required. With negative values suppressed, aggregation is much more complicated, effectively requiring that users first impute the suppressed data. Publishing the negative values would also improve users' ability to impute positive values, as the negative values combined with the published MOEs would provide meaningful information about the range of values the true counts are likely to have.

Fundamentally, if users have access to all outputs from SafeTab-P (positive and negative counts), then they have the maximum safely publishable information to use in their analyses, and they have the option to suppress negative counts if they choose. Why preemptively take away this information?

"Race alone" counts that exceed "race alone or in combination" counts

In the real world, a "race alone or in combination" count should always be equal to or greater than the corresponding "race alone" count. SafeTab-P's noise infusion may produce situations where the "race alone" count exceeds the "race alone or in combination" count. Where that happens, the Bureau proposes to suppress the "race alone" count.

Given the relatively small MOEs for the DDHC-A, suppression is most likely when the "race alone" and "race alone or in combination" counts are similar. If the "race alone or in combination" is substantially higher than the "race alone" count, the probability that noise infusion will yield a "race alone" count greater than or equal to the "race alone or in combination" count is low.

This suppression rule is applied on a line-by-line basis, and complementary suppression is not applied. This may lead to some surprising results. Consider the theoretical example in Table 2.[1]

Table 2. Theoretical noise-infused counts for Mexican alone and Mexican alone or in combination in South Dakota.

South Dakota	Mexican alone	Mexican alone or in combination	MOE
Total population:	650	670	8.5
Male:	310	330	6
Under 18 years	140	145	3
18 to 44 years	110	113	3
45 to 64 years	40	48	3
65 years and over	20	24	3
Female:	340	340	6
Under 18 years	150	151	3
18 to 44 years	120	121	3
45 to 64 years	45	46	3
65 years and over	x	22	3

This table shows the noise-infused total, sex, and sex by age counts for Mexican alone and Mexican alone or in combination detailed group in South Dakota. Based on the adaptive design of the disclosure avoidance system, the noise infused total population count for Mexican alone meets the threshold for the four-category age table. The noise infused "Mexican alone, female, age 65 years and over" count is suppressed because it was larger than the noise infused "Mexican alone or in combination, female, age 65 years and over" count.

In a traditional suppression routine, the female and the total population counts for Mexican alone would also be suppressed since one of their component counts was suppressed. However, in the line-by-line suppression approach, the noise-infused female and total population counts are reported in the data, so a user could just subtract the sum of the three other age categories for "Mexican alone, female" (150+120+45 = 315) from the female count (340) to get the suppressed count for age 65 years and over (25).

The "Mexican alone, female, age 65 years and over" count (25) is within the MOE (3) of the noise-infused count for "Mexican alone or in combination, female, age 65 years and over" (22). This is to be expected based on the design of SafeTab-P and should be viewed as unsurprising given the reported uncertainty introduced by the noise infusion routine.

[1] The Bureau's Proof of Concept Examples spreadsheet (Nation & State Detailed Groups tab) includes an example of this suppression type. The count of "Mexican alone, females, age 85 and older "is suppressed because it exceeded the count of "Mexican alone or in combination, females, age 85 and older" in Arizona.

The suppression in these cases is not intended to protect the privacy of respondents. It serves only to mask the messiness of noise-infused data. By suppressing the data in these situations, the Census Bureau is (1) hiding useful privacy-protected data that could safely be published and (2) placing a new burden on the end user to account for suppression in their analysis, either by backing out the suppressed count when possible (as in Table 2), by imputing the data where exact computation is not possible, or by using data for a detailed "alone or in combination" group to avoid suppressed data. As discussed above for the negative-value case, suppression also unnecessarily complicates users' ability to aggregate counts for multiple categories or geographic units.

The Census Bureau chose to use a noise infusion algorithm for disclosure avoidance, and this algorithm can generate counts that are "demographically unreasonable." This will be problematic for many users, but sweeping such cases under the rug with suppression does not make them go away. It only further complicates things. The Census Bureau should instead publish the counts and then clearly explain how they come about, adding guidance for how to address problems that may arise.

Thinking ahead to the DDHC-B product

As far as we know, the Census Bureau will use the same algorithm and suppression rules for the DDHC-B product (detailed race, ethnicity, and American Indian and Alaska Native tribal and village group counts for household type and tenure). If true, data suppression, particularly due to negative counts, will be more common in the DDHC-B product. There are many fewer households than people in the United States. Thus, confidential counts for households will be smaller and more of them will be close to zero. If the same privacy-loss budget and allocation is used for DDHC-B, the algorithm will produce more negative counts, which will be suppressed in the published data.

Suppression due to noise-infused detailed "race alone" counts exceeding noise-infused "race alone or in combination" counts is also expected to be more pervasive in the DDHC-B product. Since there are fewer households than persons, the household counts will be smaller and closer together in absolute terms. With the same privacy-loss budget, this increases the likelihood that a "race alone" count exceeds the "race alone or in combination" count.

Use of FOIA to force the release of suppressed data

The noise-infused counts are differentially private and meet the Bureau's requirements for protecting confidentiality. There is no apparent legal requirement for the Bureau not to publish the noise-infused counts, only its own internal decision that "demographically unreasonable" data should be suppressed. If we were policy makers or a civil rights organization seeking more complete data, we would immediately file a FOIA request to have the suppressed data released to the public, and we expect the request would need to be granted. Publishing the unreasonable data from the outset would immediately eliminate any potential future costs for groups who would make such requests and for the Bureau to respond to them.

Improve messaging about data aggregation

The SafeTab-P algorithm infuses noise with a pre-specified margin of error directly into DDHC-A counts, and aggregating counts (e.g., for two or more geographic units or for different age groups) yields increased variability. Thus, throughout the Proof of Concept document (pp. 6-7, 15, 19) and the FAQ (p. 3), the Bureau recommends that users "should use the published statistic they are interested in, when

available, rather than aggregating data" (Proof of Concept, pp. 6-7). However, there are many settings where aggregating counts will be necessary and/or beneficial, and as we understand, it should generally result in counts with MOEs that are smaller relative to the absolute count, so it is not clear where or why aggregation should be avoided.

However, users may not be able to directly use the statistic of interest as published due to the adaptive design aspect of the disclosure avoidance system. Different geographic units may have different sets of categories, thus necessitating aggregation over categories to compare data for two or more geographic units, For example, the Proof of Concept Examples spreadsheet (Sub-State Detailed Groups tab) includes sex by age counts for Mexicans in Amador and San Bernardino Counties. San Bernardino includes 23 age categories, but Amador only has four. To directly compare these counts, we must aggregate San Bernardino's age categories to match Amador's age categories. Users may read the Bureau's recommendations and worry about whether they should compute such aggregations.

Instead of cautioning against aggregation, the Bureau should clearly explain the costs and benefits of aggregating data and provide the formulas for computing new margins of error for aggregated counts. A clear explanation plus the formulas empowers users to decide whether they should aggregate instead of simply heeding the Bureau's cautionary statements against aggregation.

The Bureau already provides such guidance for the American Community Survey. The Compass Guides provide appendices describing uncertainty in the ACS and how to calculate margins of error for derived statistics, including statistics based on aggregation. We recommend that the Bureau develop a similar guide for the DDHC products, describing what happens when users aggregate data and providing instructions to compute new margins of error for aggregated statistics.

Closing thoughts

With no explanation given for why the Bureau is proposing to suppress "demographically unreasonable" data, we are left to suppose that this is a response to the critiques they have faced about impossible and implausible results in the 2020 Redistricting Data (e.g., blocks with kids and no adults, blocks with occupied housing units and no people). Instead of publishing such results in future products, it seems the Bureau has decided to use suppression to avoid such criticism. But suppression only obscures unreasonable results without improving them while also introducing a new set of problems that will in turn invite—and deserve—further criticism.

The Bureau has continually criticized suppression as an insufficient method to protect respondent confidentiality. Yet, they now turn to suppression to hide the problematically unreasonable results of their chosen alternative approach. If the Bureau is truly committed to differential privacy, they should embrace the outputs and defend them. The Bureau should develop a communication and education plan that defends their decisions and teaches data users about the algorithm.

Finally, the messaging related to data aggregation should be more explanatory and less cautionary. The Bureau's ACS documentation provides a blueprint for how to do this. We strongly encourage the Bureau to develop similar documentation for the DDHC products.

3. Chloe Rinehart, Apex for Youth

Wed 2/8/2023 4:02 PM

Hi there,

I just attended the AANHPI Proof of Concept Webinar, and wanted to provide the feedback that some aspects of the current specifications would negatively impact my work, namely, the downside of noise aggregation impacting the ability to look at data by a city neighborhood-level, for example.

It's important to my work and many others in NYC to be able to look at characteristics of AANHPI communities by neighborhood, but it sounds like the aggregation of census tracts would aggregate noise in the population estimates to a degree that would make neighborhood-level aggregates unreliable.

It would be necessary to receive guidance from the Census Bureau on how to account for this noise aggregation in this and other contexts, for example, also in combining age groups/ categories.

thanks!

Chloe

--

Chloe Rinehart
she/her/hers
Program Data Analysis Manager

4. Amy Harth, City of Minneapolis

Mon 2/13/2023 9:29 AM

Thanks so much for making such a clear presentation on the Detailed DHC-A and what to expect. I truly appreciate the opportunity to offer feedback:

1. I am incredibly happy that data will be offered at the tract and place level. I was hoping for block group so I could do custom aggregations to Minneapolis, MN neighborhoods. However, it's understandable that's not possible. I have a workaround where I use Tableau to overlay neighborhood boundaries on top of a map using tracts so people can get a general idea. I can't do neighborhood level bar graphs without block groups, but having tract is far better than not having any granular data. I suspect that even if you provided block group, many of the numbers would be hard to work with as we're now talking very small populations with the breakouts provided in the DHC-A, so if I were looking at the actual data, tract might be preferable in this case, regardless.

   If you want to see an example of how I use tract in my dashboards, check out the Race/Ethnicity page of this dashboard. Contrast that with the Poverty Ratio tab where I'm able to aggregate to neighborhood:
   https://www.minneapolismn.gov/government/government-data/datasource/poverty-dashboard/

2. I'm puzzled by the guidance that we can't do custom aggregations. Yes, it inflates the margin of error. However, if I'm understanding correctly, your margins of error are trivial compared to the ACS data I typically work with. Also, ACS uses a 90% CI and you're using 95%. Therefore, I honestly don't see where you're coming from with that advice.

   Could you provide guidance for those of us who care a lot about doing the aggregations we need and can accept having a somewhat inflated MOE? As long as I can calculate and present the MOE and Coefficient of Variation, along with the estimate, my users will be quite used to seeing things that way. See the dashboard linked to above that includes MOEs as well as a reliability estimate that's based on the calculated Coefficient of Variation.

   If you don't provide such guidance, people who don't understand MOEs and statistics are simply going to grab their calculators and do it anyway, even if I advise them not to when I create dashboards. People who work with youth want to see age breakouts as well as total youth under 18, for example. So I can provide the breakouts you offer, but they're just going to turn around and grab their calculators when it doesn't meet their needs. Better if I provide correctly calculated numbers and guidance on interpretation as I did with the ACS dashboards.

3. Could you provide additional tables that do not break the data out by sex when age is a breakout? I've never had a single person ask me to break out by sex, only age. (Note that this will change if gender categories are modernized as there are definitely folks who want that data). I wish the ACS would provide age tables that don't disaggregate sex as well. There are so many age categories that further dividing by sex is really inflating my MOE,

because I always have to aggregate the two sexes together, on top of all the other aggregations I do.

No doubt I've pushed beyond the limits of what the ACS staff would advise. However, people really care about age for many reasons, so I can't simply throw up my hands and say I'm not pleased with the wildly inflated MOEs I'm introducing. Providing programming for youth is important, for example, so I created an entire dashboard around youth: https://www.minneapolismn.gov/government/government-data/datasource/youth-dashboard/

Again, many thank yous for engaging your user base. I'm continually impressed and pleased with the work you all do, and it makes a huge difference to me and the people who consume the dashboards I create.

Amy Harth
Data Scientist – Enterprise Data + Solutions: Data + Analytics Services Group

City of Minneapolis – Information Technology

5. Orange County Asian and Pacific Islander Community Alliance

Mon 2/27/2023 2:20 PM

Dear Census Administrators:

We at the Orange County Asian & Pacific Islander Community Islander (OCAPICA) welcome the opportunity to submit comments on the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) prototype released on January 31, 2023. The Census Bureau's responsiveness to the community-based feedback that resulted in prioritizing the release of population count data for detailed race and ethnic groups through the Detailed DHC-A is commended. We look forward to the Census Bureau's responses to the feedback generated by this prototype release.

The importance of the 2020 Census data for marginalized communities, such as immigrants, those with limited English proficiency, and communities of color, cannot be overemphasized. The effort to reach every American through outreach and partnerships with community-based organizations as well as the language and cultural support for those efforts resulted in a level of awareness, trust, and expectation that the 2020 Census will continue to represent the gold standard for demographic data for our nation. For many communities and ethnic groups, the 2020 Census represents one of the few opportunities to be seen in government data and is viewed as more accurate than the American Community Survey (ACS) due to the major investment in outreach to build awareness and encourage participation in the 2020 Census by the Census Bureau, state and local governments, philanthropy, and community-based organizations.

In the documents released as part of the Detailed DHC-A prototype, the Census Bureau cautions against aggregating the Detailed DHC-A data beyond the geographies in the published tables because the method that privacy-ensuring noise has been infused into the data results in variance that is additive in nature and increases as more data points are aggregated. Unfortunately, there are a large number of use cases requiring aggregating data that communities depend on for advocacy, planning, and evaluating their programs and policies. For example, the Census Bureau does not tabulate for legislative district boundaries below the state level. Advocates rely on Census tract data to estimate the Asian ethnic group populations by City Council District to inform elected officials about the demographics of the districts they represent. Community organizations rely on aggregating Census tract data to define the neighborhoods they serve for planning and grant-writing processes.

It is essential that the Census Bureau recognize that once the data is released, data users will have the expectation to be able to use the 2020 Census data in similar ways as past decennial census data. In response to the Detailed DHC-A prototype documents, we at OCAPICA make the following recommendations:

1. The Census Bureau should provide specific guidance on how to compare 2010 vs 2020 population totals for groups and geographies that are available. Due to changes in how the race and ethnicity questions were asked and how they were coded, the Census Bureau recommends

caution when comparing the two censuses. We encourage the Census Bureau to provide additional guidance on what to watch out for when making these comparisons because these comparisons will inevitably happen.

1. The Census Bureau should provide clear and accessible information on the infused margin of errors due to the application of privacy protections. Ideally, the margins of error should appear directly in the data tables, rather than embedded in technical documentation.

1. There should be guidance on how to estimate the noise-infused margins of errors when users aggregate geographies or categories such as age to create totals for non-Census geographies, despite the Census Bureau's cautions against doing so.

There are many government jurisdictions for which the Census Bureau does not provide tabulations: for example, City Council Districts and County Legislative Districts. There are also neighborhoods that are larger than a Census tract that depend on Census level data to understand the changes in population and demographics. We will want clear ways to estimate the infused margins of error to gauge the accuracy of the aggregated noise-infused counts.

We recommend that the Census Bureau provide tools to easily estimate the variance and compare data points, similar to the Statistical Testing Tool the Census Bureau developed for the ACS: https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html

1. We also ask the Census Bureau to provide guidance on estimating the variance for ratios and percentages. Data users will often compare the percent share of an ethnic group in tracts or neighborhoods to help delineate service areas or the percent share of age groups to identify whether a neighborhood needs additional age-based services, such as daycare or senior centers. Understanding the reliability and accuracy of aggregated count data is essential for advocacy, planning, and evaluation.

1. We are concerned about whether the population thresholds for regional groups are low enough to provide aggregated data for communities that share some cultural or linguistic characteristics in areas where the detailed groups themselves would not meet reporting criteria. We are particularly interested if there are cases where none of the detailed groups in a regional group meet the population threshold for reporting (count of 22) in a geography, and the regional group also fails to meet the population threshold of 94, resulting in no disaggregated data for the populations that fall under that regional group for that geography.

1. We would like to know if the data in the Detailed DHC files (A, B, and S) would be used for the population estimates program and the population estimates produced by the ACS.

1. Data users will likely compare Detailed DHC-A population totals with ACS population estimates for various groups. The Census Bureau should provide guidance on how to compare the two datasets, and if and how statistical testing and comparisons should be made using

noise-infused margins of error in the Detailed DHC-A and sampling-based margins of error of the ACS.

1. Given the increased importance of the regional groups under the Race and Ethnicity Coding system, we have the following questions about the regional group definitions as outlined in the 2020 Census State Redistricting Data (Public Law 94-171) Summary File technical documentation (found here: https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf)

* Why is Hmong considered East Asian? Much of the immigration/refugee patterns for the Hmong community in the US are of Southeast Asian origin. The refugee experiences of the Hmong community have more commonality with Southeast Asians than with East Asians.

* Will some of the codes under Other Asian be included in regional groups? For example, Indo-Chinese typically refers to Vietnamese and would be under the Southeast Asian regional group. Malay likewise is typically associated with the Southeast Asian regional group. Urdu is typically associated with South Asian. Timorese is likely Southeast Asian.

* How are Guyanese Indian under the Some Other Race category data tabulated? Are these individuals who identify as Indo-Guyanese? The Indo-Caribbean community is part of the South Asian diaspora and could be also tabulated as Asian in combination or South Asian in combination, along with their current Caribbean regional group.

* How are multiracial/multiethnic codes 8500-8999 tabulated, specifically the Amerasian and Eurasian categories? Are those two codes in particular included as Asian in combination?

1. For Limitation 3B, where noise-infused count totals for Alone counts are larger than noise-infused counts for alone or in any combination are suppressed, the Census Bureau must clearly indicate to data users in the Alone tables where the data is suppressed that the Alone or in combination count is available and should be used.

Thank you,

Julie Vo
Policy Director

ocapica
orange county asian and pacific islander community alliance

6. Southeast Asian Resource Action Center et al.

Thu 3/2/2023 10:09 AM

Dear Nicholas, Rachel, and Cynthia,

Thank you for the opportunity to comment on the 2020 Census Detailed Demographic and Housing Characteristics File A. Please find attached a Hmong-led letter signed by 57 organizations in response to the coding of Hmong under the East Asian regional group.

While we understand that the Bureau has engaged in consultation with experts, we are concerned that there have been little direct conversation with Hmong and Southeast Asian leaders and organizations. We are further concerned that this coding decision invalidates our communities' long-held identity in the United States.

We hope that the Bureau strongly considers our recommendations to code Hmong as Southeast Asian and meet with Hmong community leaders and organizations. If you have any questions or comments, please feel free to reach out.

--

Best,

Kham S. Moua

Pronouns: (he/him/his)

National Deputy Director

Southeast Asia Resource Action Center

March 2, 2023

Nicholas A. Jones
Director and Senior Advisor of Race and Ethnic Research and Outreach, Population Division
U.S. Census Bureau

Rachel Marks
Chief, Racial Statistics Branch, Population Division
U.S. Census Bureau

Cynthia Davis Hollingsworth
Program Manager for 2020 Census Data Products and Dissemination
U.S. Census Bureau

Re: Feedback on Hmong Coding in the Detailed DHC-A Proof of Concept

Dear Mr. Jones, Ms. Marks, and Ms. Hollingsworth:

We, the undersigned 57 Hmong American; Southeast Asian American; and Asian American, Native Hawai'ian, and Pacific Islander organizations write to express our concern regarding the inclusion of Hmong under the East Asian regional group, rather than the Southeast Asian regional group, in the 2020 Census Data Products, including the upcoming Demographic and Housing Characteristics products and last year's 2020 Census State Redistricting File (Public Law 94-171) Summary File. We urge you to recategorize Hmong as Southeast Asian and, moving forward, improve how the U.S. Census Bureau conducts outreach to diverse communities to accurately assess and represent how they self-identify.

As Hmong and Southeast Asian refugee organizations, we have advocated at all levels of government for disaggregated race and ethnicity data to ensure that our communities are seen and equitably supported by public policies. Our organizations played critical roles in ensuring our diverse communities were counted in the 2020 Census, such as providing in-language outreach and support.

Hmong Americans' origins are as refugees from Southeast Asia.

Southeast Asian American (SEAA) is not only a geographic identity, but also a political identity that comes from the shared experiences of people who came to this country as refugees from Cambodia, Laos, and Vietnam. SEAAs now number over 3 million, and the vast majority are refugees, the children of refugees, and their family members. Beginning in 1975, when the U.S. withdrew from its direct and covert military interventions, large waves of people from Cambodia, Laos, and Vietnam arrived in the U.S. as refugees fleeing war, genocide, or persecution. This community includes Hmong Americans. During the Vietnam War, the U.S. recruited Hmong people in Laos for the CIA's "Secret War," and upon the U.S.'s withdrawal from Southeast Asia, tens of thousands of Hmong people fled retaliatory persecution in Laos to resettle in the U.S. as refugees.[1] Upon enactment of the Refugee Act of 1980, the number of refugees from Southeast Asia increased exponentially, and Southeast Asian refugees made up more than half of those who sought refuge in the U.S. for the next decade. The closing of the final temporary shelter in 2004, at the Buddhist monastery at Wat Tham Krabok in Thailand, led to the final wave of Hmong refugees to resettle in the U.S. in 2004 and 2005.[2]

The U.S. Census Bureau's own data describes this history. According to the American Community Survey, nearly all foreign-born Hmong Americans migrated from Southeast Asia, which the Bureau defines as including the following countries: Myanmar, Cambodia, Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand, and Vietnam. The most recent American Community Survey estimates show that 95.6% of Hmong Americans reported Southeast Asia as their region of origin, while only 1.6% reported East Asia.[3] Improperly categorizing Hmong Americans as East Asian conceals the inequities they experience as refugees from Southeast Asia.

Today, SEAAs continue to experience disparities across multiple measures of social, health, and economic security as a result of their refugee history and their marginalization by U.S. society. Nationwide, over 1 million SEAAs are low-income, including about 460,000 SEAAs who experience poverty. All SEAA ethnic groups have lower per capita incomes than average, and Hmong Americans fare worse than all racial groups across multiple measures of income. Nearly 60% of Hmong Americans are low-income, and more than one of every four live in poverty.[4]

Around a quarter of SEAA adults, including Hmong Americans, aged 25 and older have not graduated high school, compared to 12% of all Asians. An additional 56% of Hmong American high school graduates have not continued to complete a bachelor's degree, compared to 32% of all Asians, an aggregate figure that includes many actual East Asian Americans, such as Taiwanese Americans (16%) and Chinese Americans (27%).[5]

---

[1] See: Yam, K. (2021, July 30). Hmong Americans are often obscured by model minority myth. Why Suni Lee's win means so much. NBC Asian America. https://www.nbcnews.com/news/asian-america/hmong-americans-are-often-obscured-model-minority-myt h-why-suni-n1275567 ; Rahim, S. (2022, May 16). Preserving the history of America's 'secret war' in Laos. NBC. https://www.nbcnews.com/news/world/preserving-history-americas-secret-war-laos-rcna28893

[2] Southeast Asia Resource Action Center and Asian Americans Advancing Justice – Los Angeles. (2020). Southeast Asian American Journeys: A National Snapshot of Our Communities.

[3] Analysis of the American Community Survey by AAPI Data (www.aapidata.com)

[4] Southeast Asian American Journeys: A National Snapshot of Our Communities

[5] AAPI Data. (2022 June). State of Asian Americans, Native Hawaiians, and Pacific Islanders in the United States. https://aapidata.com/wp-content/uploads/2022/06/State-AANHPIs-National-June2022.pdf

Hmong Americans' shared migration history with and similarities in outcomes as other SEAAs requires that Hmong Americans are accurately categorized as Southeast Asian by region of origin, especially if small population counts in certain areas of the country necessitate a less granular disaggregation by region of origin rather than by ethnicity. Lumping the relatively small Hmong American community – approximately 323,000 in 2020[6] – with much larger Asian American communities numbering in the millions, that originate from East Asia and have different histories and outcomes in the U.S., will obscure the disparities and unique challenges experienced by Hmong American and other Southeast Asian American communities. Without accurately identifying the Hmong American community, critical government programs that are reliant on these data sets will fail to identify the disparities and inequities this population experiences and meet its needs.

The U.S. Census Bureau should improve how it engages diverse communities to accurately assess and represent how communities self-identify.

We are concerned by the Bureau's lack of engagement of Hmong American, Southeast Asian American, and other Asian American, Native Hawai'ian, and Pacific Islander communities while developing the 2020 Census Data Products. In the Bureau's own words, the race and ethnicity questions in the 2020 Census "provide[d] ways for all respondents to self-identify their detailed identities."[7] However, the Bureau has not applied the principle of self-identification to its subsequent data products based on the results of the 2020 Census. We are further concerned that this issue has existed since at least 2021, when the Bureau released the 2020 Census State Redistricting Data (Public Law 94-171) Summary File that inaccurately categorized Hmong Americans as East Asian (see Appendix F: Hispanic Origin and Race Code List).

As members of the Hmong American, Southeast Asian American, and broader Asian American, Native Hawai'ian, and Pacific Islander communities, we request that the U.S. Census Bureau improves its engagement of our diverse communities when developing essential data sets like the Demographic and Housing Characteristics products, particularly regarding variables like race and ethnicity that are self-identified. Additionally, we request a meeting between the U.S. Census Bureau and Hmong American leaders so that the Bureau can learn more about the Hmong American community and discuss ways to better support them.

We thank you for the opportunity to comment on this critical resource for our communities, policymakers, researchers, and the public. If you would like to discuss the contents of this letter further, and to schedule a meeting with community leaders, please contact Kham S. Moua (he/him), SEARAC National Deputy Director or Anna Byon (they/them), SEARAC Director of National Policy, at anna@searac.org.

Sincerely,

Coalition of Asian American Leaders (CAAL) Freedom Inc.

---

[6] Analysis of 2020 American Community Survey Five-year Public Use Microdata Sample

[7] U.S. Census Bureau. (2021 June). 2020 Census State Redistricting Data (Public Law 94-171) Summary File.
https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf

Hmong American Partnership (HAP)
Hmong Cultural Center of Butte County (HCCBC)
Hmong Innovating Politics (HIP)
Hmong National Development (HND)
Southeast Asia Resource Action Center (SEARAC)
The Fresno Center
Additional Hmong and Southeast Asian Organizations Alliance of Rhode Island Southeast Asians for Education (ARISE) Black and Brown Womyn Power Coalition, Inc.
Cambodian American Girls Empowering Cambodian Association of Greater Philadelphia Cia Siab, Inc.
Filipino Students Association at Johns Hopkins University
Hmong American Women's Association
Khmer Community of Seattle King County
Khmer Girls in Action
ManForward
Stone Soup Fresno
Vietnamese American Roundtable
Wisconsin United Coalition of Mutual Assistance Associations, Inc. (WUCMAA)
AANHPI and Non-AANHPI Organizations
AAPI Victory Alliance
Asian & Pacific Islander American Health Forum (APIAHF)
Asian Americans Advancing Justice - AAJC
Asian Americans Advancing Justice - Atlanta
Asian and Pacific Islander American Vote (APIAVote)
Asian Pacific American Labor Alliance, AFL-CIO
Asian Real Estate Association of America
Asian Texans for Justice
Association of Asian Pacific Community Health Organizations (AAPCHO) CAP Services, Inc.
Empowering Pacific Islander Communities (EPIC)
Fresno Interdenominational Refugee Ministries (FIRM)
Hamkae Center
Hmong Wisconsin Chamber of Commerce
Japanese American Citizens League
Johns Hopkins University Inter-Asian Council
National Asian American Pacific Islander Mental Health Association
National Asian Pacific American Bar Association (NAPABA)
National Asian Pacific American Families Against Substance Abuse (NAPAFASA)
National Asian Pacific Center on Aging (NAPCA)
National Council of Asian Pacific Americans - NCAPA
National Queer Asian Pacific Islander Alliance (NQAPIA)
OCA Asian Pacific Islander American Advocates Utah Chapter
OCA Greater Chicago
OCA Utah Chapter
OCA-Asian Pacific American Advocates
OCA-Asian Pacific American Advocates, Orange County Chapter

OCA-Greater Washington DC, APA Advocates
OCA, San Francisco Chapter
Organization of Chinese Americans (Detroit Chapter)
Organization of Chinese Americans (OCA), Westchester & Hudson River Chapter
Pacific Asian Counseling Services
SHK Global Health
Storied Analytics, LLC
The Organization of Chinese American Central Virginia Chapter
VAYLA New Orleans

7. New York City Department of Planning

Thu 3/2/2023 1:18 PM

Thank you for your invitation to provide feedback on planned 2020 Census data products. Attached you will find New York City Department of City Planning's response.

Best,
Joel Alvarez

JOEL A. ALVAREZ (he/him)
Deputy Director • Population Division

DEPARTMENT OF CITY PLANNING
CITY OF NEW YORK
POPULATION DIVISION

March 2, 2023

Robert L. Santos, Director
U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233

Director Santos,

RE: 2020 Census Detailed DHC-A request for Feedback on Proof of Concept (https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/newsletters/released-today-detailed-dhc-a-proof-of-concept.html), January 31, 2023.

On behalf of the New York City Department of City Planning (DCP), I am pleased to respond to the January 31, 2023 request for feedback on the 2020 Census Detailed DHC-A Proof of Concept.

Vast amounts of 2020 Census data – beyond the basic housing and demographic characteristics included in the latest 2020 redistricting data – remain unpublished. One important upcoming publication is the Detailed DHC-A providing critical information about detailed race and Hispanic ethnic groups. The Census Bureau's current proposal for this publication, as summarized in the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) Proof of Concept document, includes geographic detail for the nation, states, counties, tracts, places, and AIANNH areas. The proposal also includes measures of data accuracy, in the form of margins of error (MOEs). Perhaps most importantly, the proposal employs an adaptive design to determine how much data detail will be published based on the population count and geographic summary level.

The Census Bureau should be commended for this proof of concept. It shows the Bureau carefully listened to data user needs. The proposal provides substantial geographic detail, it prioritizes the publication of population counts for each detailed group, it produces measures of accuracy, and it adopts a disclosure avoidance framework which emphasizes tabular output, creatively employs publication thresholds, and minimizes noise infusion while still protecting respondent privacy.

Arun Peter Lobo Ph.D., Chief Demographer
Population Division
Department of City Planning
120 Broadway – 31st Floor, New York, N.Y. 10271-0001
www.nyc.gov/population

The publication of MOEs will greatly enhance users' understanding of accuracy and increase data utility by providing a more precise appreciation of limitations. The use of population thresholds recognizes the specificity of our varied landscape and provides detail where counts suggest the detail matters. Further, the inclusion of census tracts is a significant improvement over the Bureau's initial proposal, which would have left places like New York City without information below the county level.

While DCP endorses most of the proposed Detailed DHC-A release plan, there are several points of concern which need to be addressed.

In the Limitations section of the Census Bureau's Proof of Concept document, there is a warning which reads:

> "Due to the limitations explored below, we caution data users against aggregating data. This includes creating new regional groups from detailed groups, combining lower-levels of geography to create higher-levels of geography, combining sex by age data, etc. We strongly recommend using the published noise infused counts associated with your geographic level of interest when available."

This cautionary note against data aggregations is simply not enough. Custom aggregations of data are commonplace for many census data users. They are necessitated by realities on the ground. In New York City, the creation of aggregate geographies by grouping census statistical geographies is essential as our community tries to approximate areas like police precincts, flood zones, and Community Districts. This is a critical part of understanding our diverse terrain. It is particularly important for a municipality like New York City, where we often aggregate census tracts into Neighborhood Tabulation Areas, which have great resonance in our communities.

Consequently, we recommend that the Census Bureau goes further beyond the current admonishment. As good data stewards, the Census Bureau should give users a sense of the pitfalls when aggregating data. Perhaps this guidance could come in the form of quantified comparisons between 2010 Census enumerated values contrasted against common values treated with the new 2020 Disclosure

Avoidance System for a set of test aggregations. This guidance would be incredibly helpful and avoid the data user paralysis which might follow the current blanket warning.

While addressing this lack of guidance on data aggregation is DCP's primary concern, it would be helpful if the Census Bureau provided further guidance in other realms.

Comparisons across decennial censuses are invaluable when gauging trajectories of change, but questionnaire updates complicate such comparisons. Therefore, the Bureau should advise data users on the pitfalls of comparisons across decades. Along the same lines, it is also inevitable that data users will want to compare DDHC-A data with American Community Survey (ACS) data. Accordingly, the Census Bureau should explain the limitations of such comparisons. Finally, the Bureau should provide guidance on the calculation of MOEs for derived statistics, like ratios and percents. The use of such statistics is commonplace and must be supported.

Beyond the provision of guidance on these topics, DCP encourages the Census Bureau to prominently publish MOEs in tabular output of count data. This will help avoid misunderstandings about the limitations of data accuracy.

In summary, DCP strongly encourages the Census Bureau to modify its warning against data aggregations (both of variables and geographies) and support the data user community with detailed quantitative guidance on the impact such aggregations have on data accuracy. Additionally, the Census Bureau should provide further guidance on comparisons across censuses, comparisons with the ACS, and the calculation of MOEs for derived statistics. Finally, DCP recommends that the

Census Bureau prominently publish MOEs, accompanying all count data.

Sincerely,

Arun Peter Lobo, Ph.D.
Chief Demographer, Population Division

8. Asian Americans Advancing Justice Southern California

Thu 3/2/2023 2:07 PM

Hello,

Asian Americans Advancing Justice Southern California (AJSOCAL) is grateful for the opportunity to submit comments on the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) prototype released on January 31, 2023. The Census Bureau's responsiveness to the community-based feedback that resulted in prioritizing the release of population count data for detailed race and ethnic groups through the Detailed DHC-A is appreciated. We look forward to the Census Bureau's responses to the feedback generated by this release.

The importance of the 2020 Census data for marginalized communities, such as immigrants, those with limited English proficiency, and communities of color, cannot be overemphasized. For many communities and ethnic groups, such as the Asian American, Native Hawaiian and Pacific Islander communities that AJSOCAL serves, the 2020 Census represents one of the few opportunities to be seen in government data and is viewed as more accurate than the American Community Survey (ACS) due to the major investment in outreach to encourage participation in the 2020 Census by the Census Bureau, state and local governments, philanthropy, and community-based organizations.

In response to the Detailed DHC-A prototype documents, we at AJSOCAL make the following recommendations:

1. The Census Bureau should provide specific guidance on how to compare 2010 vs 2020 population totals for groups and geographies that are available. Due to changes in how the race and ethnicity questions were asked and how they were coded, the Census Bureau recommends caution when comparing the two censuses. The Asian American population in the United States is among the fastest growing racial/ethnic groups. Researchers need to be able to make comparisons to understand patterns of the growth of communities across the US. As such, we encourage the Census Bureau to provide additional and clear guidance, particularly regarding the caveats of comparing two censuses that use different methods.
2. The Census Bureau should provide clear and accessible information on the infused margin of errors due to the application of privacy protections. Ideally, the margins of error should appear directly in the data tables, rather than embedded in technical documentation.
3. In the documents released as part of the Detailed DHC-A prototype, the Census Bureau cautions against aggregating the Detailed DHC-A data beyond the geographies in the published tables because the method that privacy-ensuring noise has been infused into the data results in variance that is additive in nature and increases as more data points are aggregated.

   However, there are several instances requiring aggregating data that communities depend on for advocacy, planning, and evaluating their programs and policies. For example, there are many government jurisdictions for which the Census Bureau does not provide tabulations such as City Council Districts, County Legislative Districts, and neighborhood council districts. There are also neighborhoods that are larger than a Census tract that depend on Census level data to understand the changes in population and demographics. The census tracts are too small by themselves and the readily available larger levels of geography are not practical for communities. For example, the City of Los Angeles has a population of nearly 4 million people and covers a large geographic area with 15 city council

districts which still each have over a quarter million in population. It is necessary to be able to have usable data to understand the neighborhoods and communities within these large jurisdictions. Additionally, AJSOCAL often aggregates census tracts to understand the characteristics of the specific neighborhoods within cities or counties where there are ethnic enclaves, such as Koreatowns, Chinatowns, Little Saigons, etc. Community organizations rely on aggregating Census tract data to define the neighborhoods they serve for planning and grant-writing processes.

Therefore, the Census Bureau should provide guidance on how to estimate the noise-infused margins of errors when users aggregate geographies or categories such as age to create totals for non-Census geographies, rather than simply cautioning against doing so. We also recommend that the Census Bureau provide tools to easily estimate the variance and compare data points, similar to the Statistical Testing Tool the Census Bureau developed for the ACS: https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html

4. We also ask the Census Bureau to provide guidance on estimating the variance for ratios and percentages. Data users will often compare the percent share of an ethnic group in tracts or neighborhoods to help delineate service areas or the percent share of age groups to identify whether a neighborhood needs additional age-based services, such as daycare or senior centers. Understanding the reliability and accuracy of aggregated count data is essential for advocacy, planning, and evaluation.

5. We are concerned about whether the population thresholds for regional groups are low enough to provide aggregated data for communities that share some cultural or linguistic characteristics in areas where the detailed groups themselves would not meet reporting criteria. We are particularly interested if there are cases where none of the detailed groups in a regional group meet the population threshold for reporting (count of 22) in a geography, and the regional group also fails to meet the population threshold of 94, resulting in no disaggregated data for the populations that fall under that regional group for that geography.

6. We would like to know if the data in the Detailed DHC files (A, B, and S) would be used for the population estimates program and the population estimates produced by the ACS.

7. Data users will likely compare Detailed DHC-A population totals with ACS population estimates for various groups. The Census Bureau should provide guidance on how to compare the two datasets, and if and how statistical testing and comparisons should be made using noise-infused margins of error in the Detailed DHC-A and sampling-based margins of error of the ACS.

8. Given the increased importance of the regional groups under the Race and Ethnicity Coding system, we have the following questions about the regional group definitions as outlined in the 2020 Census State Redistricting Data (Public Law 94-171) Summary File technical documentation (found here: https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf)

1. Why is Hmong considered East Asian? Much of the immigration/refugee patterns for the Hmong community in the US are of Southeast Asian origin. The refugee experiences of the Hmong community have more commonality with Southeast Asians than with East Asians.
2. Will some of the codes under Other Asian be included in regional groups? For example, Indo-Chinese typically refers to Vietnamese and would be under the Southeast Asian regional group. Malay likewise is typically associated with the Southeast Asian regional group. Urdu is typically associated with South Asian. Timorese is likely Southeast Asian.
3. How are Guyanese Indian under the Some Other Race category data tabulated? Are these individuals who identify as Indo-Guyanese? The Indo-Caribbean community is part of the South Asian diaspora and could be also tabulated as Asian in combination or South Asian in combination, along with their current Caribbean regional group.
4. How are multiracial/multiethnic codes 8500-8999 tabulated, specifically the Amerasian and Eurasian categories? Are those two codes in particular included as Asian in combination?

9. We recommend that the Census Bureau collaborate with advocacy groups and community-based organizations that represent and serve the above communities for guidance on how they identify their regions.

10. For Limitation 3B, where noise-infused count totals for Alone counts are larger than noise-infused counts for alone or in any combination are suppressed, the Census Bureau must clearly indicate to data users in the Alone tables where the data are suppressed that the Alone or in combination count is available and should be used.

Thank you again for the opportunity to provide feedback. We look forward to the Census Bureau's response to the input received.

June Lim (she/her/hers)
Demographic Research Project Director
Asian Americans Advancing Justice Southern California
Formerly known as Asian Americans Advancing Justice – Los Angeles

9. Asian American Federation

Thu 3/2/2023 3:12 PM

To whom it may concern,

Please find attached is the comments regarding 2020 Census Detailed DHC-A from Asian American Federation.

Best,

Linying He, Ph.D.
Associate Director of Research
Asian American Federation

In the documents released as part of the Detailed DHC-A prototype, the Census Bureau cautions against aggregating the Detailed DHC-A data beyond the geographies in the published tables because the method that privacy-ensuring noise has been infused into the data results in variance that is additive in nature and increases as more data points are aggregated. Unfortunately, there are a large number of user cases requiring aggregating data that communities depend on for advocacy, planning, and evaluating their programs and policies. For example, the Census Bureau does not tabulate for legislative district boundaries below the state level. Advocates in New York City rely on Census tract data to estimate the Asian ethnic group populations by City Council District to inform elected officials about the demographics of the districts they represent. Community organizations rely on aggregating Census tract data to define the neighborhoods they serve for planning and grant-writing processes.

It is essential that the Census Bureau recognize that once the data is released, data users will have the expectation to be able to use the 2020 Census data in similar ways as past decennial census data. In response to the Detailed DHC-A prototype documents, Asian American Federation makes the following recommendations:

1. The Census Bureau should provide specific guidance on how to compare 2010 vs 2020 population totals for groups and geographies that are available. Due to changes in how the race and ethnicity questions were asked and how they were coded, the Census Bureau recommends caution when comparing the two censuses. We encourage the Census Bureau to provide additional guidance on what to watch out for when making these comparisons because these comparisons will inevitably happen.

2. The Census Bureau should provide clear and accessible information on the infused margin of errors due to the application of privacy protections. Ideally, the margins of error should appear directly in the data tables, rather than embedded in technical documentation.

3. There should be guidance on how to estimate the noise-infused margins of errors when users aggregate geographies or categories such as age to create totals for non-Census geographies, despite the Census Bureau's cautions against doing so.

   There are many government jurisdictions for which the Census Bureau does not provide tabulations: for example, City Council Districts and County Legislative Districts. There are also neighborhoods that are larger than a Census tract that depend on Census level data to understand the changes in population and demographics. We will want clear ways to estimate the infused margins of error to gauge the accuracy of the aggregated noise-infused counts.

   We recommend that the Census Bureau provide tools to easily estimate the variance and compare data points, similar to the Statistical Testing Tool the Census Bureau developed for the ACS: https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.html

4. We also ask the Census Bureau to provide guidance on estimating the variance for ratios and percentages. Data users will often compare the percent share of an ethnic group in tracts or neighborhoods to help delineate service areas or the percent share of age groups to identify whether a neighborhood needs additional age-based services, such as daycare or senior centers. Understanding the reliability and accuracy of aggregated count data is essential for advocacy, planning, and evaluation.

5. We are concerned about whether the population thresholds for regional groups are low enough to provide aggregated data for communities that share some cultural or linguistic characteristics in areas where the detailed groups themselves would not meet reporting criteria. We are particularly interested if there are cases where none of the detailed groups in a regional group meet the population threshold for reporting (count of 22) in a geography, and the regional group also fails to meet the population threshold of 94, resulting in no disaggregated data for the populations that fall under that regional group for that geography.

6. We would like to know if the data in the Detailed DHC files (A, B, and S) would be used for the population estimates program and the population estimates produced by the ACS.

7. Data users will likely compare Detailed DHC-A population totals with ACS population estimates for various groups. The Census Bureau should provide guidance on how to compare the two datasets, and if and how statistical testing and comparisons should be made using noise-infused margins of error in the Detailed DHC-A and sampling-based margins of error of the ACS.

8. Given the increased importance of the regional groups under the Race and Ethnicity Coding system, we have the following questions about the regional group definitions as outlined in the 2020 Census State Redistricting Data (Public Law 94-171) Summary File technical documentation (found here: https://www2.census.gov/programs-surveys/decennial/2020/technical-documentation/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf)
   a. Why is Hmong considered East Asian? Much of the immigration/refugee patterns for the Hmong community in the US are of Southeast Asian origin. The refugee experiences of the Hmong community have more commonality with Southeast Asians than with East Asians.
   b. Will some of the codes under Other Asian be included in regional groups? For example, Indo-Chinese typically refers to Vietnamese and would be under the Southeast Asian regional group. Malay likewise is typically associated with the Southeast Asian regional group. Urdu is typically associated with South Asian. Timorese is likely Southeast Asian.
   c. How are Guyanese Indian under the Some Other Race category data tabulated? Are these individuals who identify as Indo-Guyanese? The Indo-Caribbean community is part of the South Asian diaspora and could be also tabulated as Asian in combination or South Asian in combination, along with their current Caribbean regional group.
   d. How are multiracial/multiethnic codes 8500-8999 tabulated, specifically the Amerasian and Eurasian categories? Are those two codes in particular included as Asian in combination?

9. For Limitation 3B, where noise-infused count totals for Alone counts are larger than noise-infused counts for alone or in any combination are suppressed, the Census Bureau must clearly indicate to data users in the Alone tables where the data is suppressed that the Alone or in combination count is available and should be used.

10. AAPI Data

Thu 3/2/2023 5:01 PM

Dear Census Bureau Staff,

Please see the attached PDF comment regarding the 2020 Census Detailed DHC-A from the team at the AAPI Data project.

Thank you for the opportunity to comment,

Howard

--

Howard Shih

Managing Director

AAPI Data

info@aapidata.com
aapidata.com

March 2, 2023

Robert L. Santos, Director
U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233

Re: Comments on 2020 Census Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) prototype Dear Director Santos:

AAPI Data welcomes the opportunity to submit comments on the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) prototype released on January 31, 2023. The Census Bureau's responsiveness to the community-based feedback that resulted in prioritizing the release of population count data for detailed race and ethnic groups through the Detailed DHC-A is commended.

The importance of the 2020 Census data for underserved and marginalized communities, such as immigrants, those with limited English proficiency, and communities of color, cannot be overemphasized. The effort to count every American through outreach and partnerships with community-based organizations as well as the language and cultural support for those efforts resulted in a level of awareness, trust, and expectation that the 2020 Census will continue to represent the gold standard for demographic data for our nation.

For many communities and ethnic groups, the 2020 Census represents one of the few opportunities to gain visibility in government data and is viewed as more accurate than the American Community Survey (ACS) due to the major investment in outreach to build awareness and encourage participation in the 2020 Census by the Census Bureau, state and local governments, philanthropy, and community-based organizations.

In the documents released as part of the Detailed DHC-A prototype, the Census Bureau cautions against aggregating the Detailed DHC-A data beyond the geographies in the published tables because the method that infused privacy-ensuring noise into the data results in variance that is additive in nature and increases as more data points are aggregated. Unfortunately, there are a large number of use cases requiring aggregating data that communities depend on for advocacy, planning, and evaluating their programs and policies. For example, the Census Bureau does not tabulate data for legislative district boundaries below the state level. Advocates in New York City rely on Census tract data to estimate the Asian ethnic group populations by City Council District to inform elected officials about the demographics of the districts they represent. Community organizations rely on aggregating Census tract data to define the neighborhoods they serve for planning and grant-writing processes.

It is essential that the Census Bureau recognize that once the data is released, data users will have the expectation to be able to use the 2020 Census data in similar ways as past decennial census data. In response to the Detailed DHC-A prototype documents, we make the following recommendations:

1) **The Census Bureau should provide specific guidance on how to compare 2010 vs 2020 population totals for groups and geographies that are available**. Due to changes in how the race and ethnicity questions were asked and how they were coded, the Census Bureau recommends caution when comparing the two censuses. We encourage the Census Bureau to provide additional guidance on what to watch out for when making these comparisons because these comparisons will inevitably happen.

2) **The Census Bureau should provide clear and accessible information on the infused margin of errors due to the application of privacy protections**. Ideally, the margins of error should appear directly in the data tables as published in tools such as data.census.gov, rather than embedded in technical documentation.

3) **There should be guidance on how to estimate the noise-infused margins of errors when users aggregate geographies or categories** such as age to create totals for non-Census geographies, despite the Census Bureau's cautions against doing so.

   There are many government jurisdictions for which the Census Bureau does not provide tabulations. For example, City Council Districts and County Legislative Districts. There are also neighborhoods that are larger than a Census tract that depend on Census level data to understand the changes in population and demographics. We will want clear ways to estimate the infused margins of error to gauge the accuracy of the aggregated noise-infused counts.

   We recommend that the Census Bureau provide tools to easily estimate the variance and compare data points, similar to the Statistical Testing Tool the Census Bureau developed for the ACS: https://www.census.gov/programs-surveys/acs/guidance/statistical-testing-tool.ht ml

4) **We also ask the Census Bureau to provide guidance on estimating the variance for ratios and percentages**. Data users will often compare the percent share of an ethnic group in tracts or neighborhoods to help delineate service areas or the percent share of age groups to identify whether a neighborhood needs additional age-based services, such as daycare or senior centers. Understanding the reliability and accuracy of aggregated count data is essential for advocacy, planning, and evaluation.

5) **We are concerned about whether the population thresholds for regional groups are low enough to provide aggregated data** for communities that share some cultural or linguistic characteristics in areas where the detailed groups themselves would not meet reporting criteria. We are particularly interested if there are cases where none of the detailed groups in a regional group meet the population threshold for reporting (count of 22) in a geography, AND the regional group also fails to meet the population threshold of 94, resulting in no disaggregated data for the populations that fall under that regional group for that geography.

6) We would like to know if the data in the Detailed DHC files (A, B, and S) would be used for the **population estimates program and the population estimates produced** out of the American Community Survey.

7) **Data users will likely compare Detailed DHC-A population totals with ACS population estimates for various groups**. The Census Bureau should provide guidance on how to compare the two datasets, and if and how statistical testing and comparisons should be made between noise-infused margins of error in the Detailed DHC-A and sampling-based margins of error of the ACS.

8) **For Limitation 3B in the Detailed Demographic and Housing Characteristics File A (Detailed DHC-A) Proof of Concept documentWe** , where noise-infused count totals for Alone counts are larger than noise-infused counts for alone or in any combination are suppressed, the **Census Bureau must clearly indicate to data users in the Alone tables where the data is suppressed that the Alone or in combination count is available and should be used.**

9) Given the increased importance of the regional groups under the Race and Ethnicity Coding system, we have the following recommendations about the regional group definitions as outlined in the 2020 Census State Redistricting Data (Public Law 94-171) Summary File technical documentation (found here: https://www2.census.gov/programs-surveys/decennial/2020/technical-documenta tion/complete-tech-docs/summary-file/2020Census_PL94_171Redistricting_StatesTechDoc_English.pdf)
    a) **Hmong should be classified as Southeast Asian, and not East Asian**. In our own analysis of the 2017-2021 ACS Five-Year Public Use Microdata Sample, we found that over 95% of the foreign-born Hmong population immigrated from Southeast Asia. Much of the immigration/refugee patterns for the Hmong community in the US are of Southeast Asian origin. The refugee experiences of the Hmong community have more commonality with Southeast Asians than with East Asians.
    b) **There are other errors in regional classification pertaining to the Other Asian category**. For example, Indo-Chinese is an archaic term that refers to Vietnamese and should be classified as Southeast Asian. Malay are Southeast Asian. Urdu speakers are South Asian. Timorese are Southeast Asian.

c) **Need more clarity on how Guyanese Indians are tabulated**. Are these individuals who identify as Indo-Guyanese? The Indo-Caribbean community is part of the South Asian diaspora and could be also tabulated as Asian in combination or South Asian in combination, along with their current Caribbean regional group.

d) **Need more clarity on how multiracial/multiethnic codes 8500-8999 are tabulated, specifically the Amerasian and Eurasian categories**. These two codes in particular need to be tabulated as Asians in combination.

10) **These errors point to the pressing need for the Census Bureau to use scientific and transparent methods in the planning, analysis, and decision making processes regarding ethnic and regional classification**. We recommend that the Census Bureau have frequent, broad, and transparent discussions with data users, community-based researchers, community leaders, and community organizations regarding all stages of the data collection process. We ask for frequency because immigration patterns especially among Asian Americans, Native Hawaiians, and Pacific Islanders are constantly evolving, with new groups emerging, such as the Burmese and Bhutanese refugee populations in the late 2000's and early 2010's, and with new geographies due to the dispersion of immigrant populations to all parts of the United States.

Classification of detailed origin groups into regional classifications should also use survey methods such as found in the 2016 National Asian American Survey (and analyzed in the 2020 article "Who Counts as Asian?" by Jennifer Lee and Karthick Ramakrishnan in Ethnic and Racial Studies 43:10), where racial and ethnic assignment/classification is a task that engages not only research experts in demography, history, and the social sciences, but also everyday residents including those who self-identify with a particular racial, detailed origin, or region group.

Finally, we ask for broad outreach efforts because the categories of Asian American and Pacific Islander cover a broad range of communities, nationalities, and ethnicities, each with their own sense of identity and needs. And we ask for transparency in order to understand decision-making on how self-identified data is processed and reported to ensure that the principles of self-identification are maintained throughout the data life cycle.

We look forward to the Census Bureau's responses to all the feedback generated by this prototype release and continued partnership on improving data equity and accessibility. Please email us info@aapidata.com with any publications or events that respond to the comments and concerns raised in our comment and those of others.

Sincerely,

Karthick Ramakrishnan
Founder and Co-Director

Janelle Wong
Co-Director

Howard Shih
Managing Director

Ryan Vinh
Data and Policy Manager

11. Los Angeles County Department of Public Health

Thu 3/2/2023 8:07 PM

Hello,

Please find feedback from the Los Angeles County Department of Public Health on the Detailed DHC-A Proof of Concept. Thank you for this opportunity to provide our input.

Sincerely,

Megha D. Shah, MD, MPH, MS (she/her)
Chief, Population Health Assessment Unit
Office of Health Assessment and Epidemiology
Los Angeles County Department of Public Health

COUNTY OF LOS ANGELES
Public Health

BARBARA FERRER, Ph.D., M.P.H., M.Ed.
Director

MUNTU DAVIS, M.D., M.P.H.
County Health Officer

MEGAN McCLAIRE, M.S.P.H.
Chief Deputy Director

313 North Figueroa Street, Suite 806
Los Angeles, CA 90012
TEL (213) 288-8117 • FAX (213) 975-1273

RASHMI SHETGIRI, M.D., M.S.H.S., M.S.C.S.
Director, Office of Health Assessment and Epidemiology
Acting, Chief Science Officer
313 North Figueroa Street, Room 127
Los Angeles, California 90012
TEL (213) 288-7755• FAX (213) 250-2594

BOARD OF SUPERVISORS

Hilda L. Solis
First District
Holly J. Mitchell
Second District
Lindsey P. Horvath
Third District
Janice Hahn
Fourth District
Kathryn Barger
Fifth District

www.publichealth.lacounty.gov
March 2, 2023

Robert L. Santos
Director
U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233

Dear Mr. Santos,

I am writing to request that the U.S. Census Bureau (USCB) consider a smaller population threshold as criteria for proposed suppression levels and that suppression levels be based on total population size (regardless of geography type) in the forthcoming 2020 Census Detailed Demographic Housing Characteristics (Detailed DHC-A) file.

The Los Angeles County (LAC) Department of Public Health (DPH) serves one of the largest and most diverse counties in the nation. LAC-DPH relies on detailed population estimates at county and sub-county geographies (including detailed race, ethnicity, and population counts by sex and detailed age) to inform our programmatic, evaluation, and outreach efforts and to reduce health inequities in our communities. This information is especially important to us in working with certain communities with relatively small populations but who are disproportionately affected by adverse health outcomes, such as our county's Native Hawaiian and Pacific Islander (NHPI), Asian, and American Indian and Alaska Native groups and subgroups (priority populations for our department). Our department also relies heavily on census tract population data to inform and prioritize local efforts.

As an example, the calculation of age-adjusted mortality and morbidity rates are an invaluable tool used by LAC-DPH in understanding the differential impact of certain health conditions among various groups and geographic areas in our county. To generate these rates, we require population denominators with information available for 23 distinct age groups. We are concerned that the proposed suppression levels

for the Detailed DHC-A would jeopardize our ability to produce age-adjusted rate calculations for all groups of interest in our county, including for NHPI and Asian subgroups. LAC has the largest population of NHPI individuals outside of Hawaii and one of the largest Asian populations of any county in the US. With the proposed population thresholds, population counts for many NHPI and some Asian detailed groups would be released with only four age groups. Consequently, we would no longer be able to present age-adjusted rates for preventable causes of death or other conditions, such as COVID-19, for NHPI and Asian subgroups as we have in the past when using the 2010 Summary File 2 (which used a suppression threshold of 100 for all geographies). While we acknowledge that the threshold used in the 2010 Summary File 2 may have been too low to sufficiently protect respondent confidentiality, we are concerned that the current proposed threshold of 20,000 is much too high and will negatively impact our ability to provide mortality and morbidity rate information for all communities of interest in our county. Furthermore, the current proposed threshold will render the calculation of rates and statistics at subcounty geographies impossible and will allow for rates to be reported for only broad race group categories.

In addition, over the past decade, California has passed legislation that requires Departments of Public Health to collect more detailed race and ethnicity data. Our analysis finds that based on 2020 5-year American Community Survey estimates for these mandated groups, we would be unable to report rates for all 5 NHPI groups (Table 1) and 7 of 16 Asian groups (Table 2). If the threshold was lowered to 3,000 persons, this would provide the ability to report on an additional 11 groups, a substantial improvement.

Finally, I am concerned about the proposal to use different population thresholds for different geographic levels in the forthcoming Detailed DHC-A file. LAC has a larger population than 40 states. By basing the population threshold on geography type, LAC would be unfairly compared to other counties with smaller populations (e.g., Alpine County in California, which has a population of 1,344).

In summary, I urge the USCB to release the forthcoming Detailed DHC-A with a lower proposed suppression threshold of 3,000 or less for counties, places, and census tracts for the 23 age and sex tabulations and that suppression levels be based on total populations sizes regardless of geography type.

Sincerely

Rashmi Shetgiri, M.D., M.S.H.S., M.S.C.S.
Director, Office of Health Assessment and Epidemiology
Acting, Chief Science Officer
Chief Science Office

Table 1: Native Hawaiian/Pacific Islander (NHPI) Groups* and Proposed Age Categories for which Detailed DHC-A will be available in LAC

Native Hawaiian and Other Pacific Islander	2020 LAC Population	Proposed 2020: Age Category
Native Hawaiian	4,509	4
Samoan	8,605	9
Tongan	3,005	4
Chamorro	3,289	4
Fijian	1,976	4

Population source: 2020 5-year American Community Survey (ACS) estimates
**The following California legislation requires data collection for the above listed NHPI groups:*

AB 1726 Section 8310.7 (7/1/2022)
https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201520160AB1726

Section 4302 ACA
https://aspe.hhs.gov/reports/hhs-implementation-guidance-data-collection-standards-race-ethnicity-sex-primary-language-disability-0

State of CA Health and Human Services Morbidity Reporting (use for COVID-19 only)
https://www.cdph.ca.gov/CDPH%20Document%20Library/ControlledForms/cdph110d.pdf)

Table 2: Asian Detailed Groups* and proposed age categories for which Detailed DHC-A will be available in LAC

Asian	2020 LAC Population	Proposed 2020: Age Category
Asian Indian	103,124	23
Bangladeshi	5,080	9
Cambodian	31,551	23
Chinese, except Taiwanese	445,822	23
Filipino	339,427	23
Hmong	1,109	4
Indonesian	9,341	9
Japanese	96,810	23
Korean	207,483	23
Laotian	3,223	4
Malaysian	767	Total only
Pakistani	11,797	9
Sri Lankan	4,108	4
Taiwanese	29,400	23
Thai	24,670	23
Vietnamese	99,342	23

Population source: 2020 5-year American Community Survey (ACS) estimates
**The following California legislation requires data collection for the above listed Asian groups:*

AB 1726 Section 8310.5 (1/1/2012
https://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=GOV&sectionNum=8310.5

AB 1726 Section 8310.7 (7/1/2022)
https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201520160AB1726

Section 4302 ACA https://aspe.hhs.gov/reports/hhs-implementation-guidance-data-collection-standards-race-ethnicity-sex-primary-language-disability-0

State of CA Health and Human Services Morbidity Reporting (use for COVID-19 only)
https://www.cdph.ca.gov/CDPH%20Document%20Library/ControlledForms/cdph110d.pdf)

12. Lisa Watkins-Victorino, Office of Hawaiian Affairs

Thu 3/2/2023 9:26 PM

Mahalo nui loa (thank you very much) for the opportunity to provide feedback to the Census Bureau on the Demographic and Housing Characteristics A (DHC -A) Proof of Concept data file. The Office of Hawaiian Affairs, as well as other Native Hawaiian serving organizations, use the Census data to inform advocacy, research, community engagement, and programmatic funding. Hawai'i has the highest diversity score in the nation, so we know the importance of data to address equity in education, health, housing, and economic stability. As the indigenous people of Hawai'i, it is imperative that we are able to access Native Hawaiian data disaggregated from Pacific Islander data. We have been able to do so with the 2010 Decennial Census SF2 file.

A review of proof of concept and related documents, webinar, FAQs, and fact sheet on adaptive design raises concern on our continuing ability to rely on the Census data. The documentation and the webinar indicate that this Census will provide 370 detailed races and 1,200 detailed categories for Native American and Alaska Native tribes; however, there is no codebook or detailed list provided. The Detailed Summary Metrics spreadsheet uses meta levels such as AIAN Detailed Races, Asian Detailed Races, and NHPI Detailed Races Alone and Alone in Combination. Again, there is no list of the specific detailed races included. The webinar makes note to review the Redistricting file technical documentation to get an idea of the possible detailed categories to be included. A review of that file, tables P1 – P4, indicates there are 288 race category combinations. None of which disaggregate Native Hawaiian from Pacific Islanders that provide for Native Hawaiian alone or Native Hawaiian alone or in combination categories.

In 2021 the Census Bureau released the Census Data Product Planning Crosswalk. This was the first file where we were able to see the race data. The only combinations available were in keeping with the OMB categories and Native Hawaiian was not disaggregated. The question was raised as to where was the detailed race information for not only Native Hawaiians but specific Pacific Island categories. The response was that the detailed race data would be available in the final DHC file. Given, we are now reviewing Proof of Concept for the DHC – A file and the same question was posed in the webinar, it is concerning there is still no information to give us confidence that the data needed and provided in the 2010 Census will continue to be available.

Currently, the impact is difficult to ascertain; however, if there is no disaggregation then the data will have limited if any, utility. Regarding geography, the SafeTab-P privacy algorithm will make it more difficult to create geographies for neighborhood areas or pull island-level data for our less populated islands since aggregation introduces more 'noise'.

Mahalo nui loa,
Lisa Watkins-Victorino
Research Director, Office of Hawaiian Affairs

13. Asian Americans Advancing Justice

Thu 3/2/2023 11:32 PM

Please find attached comments re: 2020 Census Detailed DHC-A from Asian Americans Advancing Justice | AAJC.

Terry

Terry Ao Minnis
Senior Director of Census and Voting Programs

March 2, 2022
2020DAS@census.gov

Re: 2020 Census Detailed DHC-A

Asian Americans Advancing Justice | AAJC (Advancing Justice | AAJC) is a national non-profit, non-partisan organization founded in 1991. For over thirty years, we have served as the leading Asian American voice on civil rights issues in our nation's capital. Our mission is to advance civil and human rights for Asian Americans and to build and promote a fair and equitable society for all. Advancing Justice | AAJC considers data collection and reporting to be the backbone of its mission.

Over the decades, we have worked to eliminate the barriers that have historically resulted in undercounting and underreporting Asian Americans and NHPIs in federal data collection and analysis efforts, particularly in the decennial census count. Our permanent census program monitors census policy and educates policy makers—including through testifying at Congressional hearings. We conduct community outreach and education on the surveys conducted by the Census Bureau, including running nationwide Asian American-focused campaigns for Census 2000, Census 2010, and Census 2020. Advancing Justice | AAJC has also served as a member of numerous advisory committees to the Census Bureau since 2000. Most recently, we served on the National Advisory Committee on Racial, Ethnic and Other Populations, completing our second three-year term through August 2019. Additionally, Advancing Justice | AAJC currently co-chairs the Leadership Conference on Civil and Human Rights' Census Task Force and serves as a co-coordinator of the Census Count campaign.

Advancing Justice | AAJC considers a fair and accurate census and comprehensive ACS among the most significant civil rights issues facing the country today. Our wide-ranging efforts to promote civic engagement, forge strong and safe communities, and create an inclusive society are guided significantly by objective, inclusive data on America's diverse communities and populations. We appreciate the importance of fact-based analyses and the need for disaggregated, detailed data on our community for purposes of identifying disparate access and outcomes and devising effective solutions. To that end, we offer the following feedback on the Detailed DHC-A Proof of Concept.

Overview

Detailed data are particularly critical for Asian Americans, who are among our nation's fastest growing and most diverse racial groups.[1] Often viewed as homogenous, these communities include more than 30 detailed subgroups that can differ dramatically across key social and economic indicators.[2] While Indian Americans have an average poverty rate of 6%, Mongolian Americans and Burmese Americans have a

[1] https://www.pewresearch.org/fact-tank/2021/04/29/key-facts-about-asian-origin-groups-in-the-u-s/.
[2] https://www1.nyc.gov/assets/immigrants/downloads/pdf/Fact-Sheet-NYCs-API-Immigrant-Population.pdf.

poverty rate of 25%.[3] Roughly 75% of Taiwanese Americans hold a bachelor's degree, yet only 14% of Laotian Americans do.[4] Another example can be found in health disparities. A study showed that "19.4% of Asian adults compared to 12.9% of whites report[ed] being without a usual source of health care, with Cambodian and Vietnamese [Americans] … three times more likely to skip doctor visits due to cost compared to all Asian [Americans] or U.S. residents."[5] The study further found that U.S.-born Vietnamese

American women represent one of the highest risk groups for breast cancer; they are four times more likely to die of breast cancer than any other Asian American group. Moreover, Korean American children are four times more likely to have no health insurance as compared to others. Finally, disaggregating data on the prevalence of smoking in New York City showed that while the prevalence in smoking was lower overall in Asian Americans compared to whites (14.1% vs. 18.6%), the actual prevalence of smoking was much higher for some Asian American subgroups, such as 35.5% in Korean Americans.[6] And while Japanese Americans, Filipino Americans, and

Indian Americans all have English language proficiency at or above 80%, only 36% of Bhutanese Americans speak English proficiently.[7]

Accuracy and Usability

Having accurate data is necessary, but having usable data is equally important. We understand that the adaptive design allows for more control of noise injection and for the advance determination of all margins of error that are met 95% of the time. However, we also know that there are limitations to utilizing this methodology.

The population thresholds for detailed group reporting for 2020 is an improvement as compared to 2010. Detailed groups with populations of 22 or more and regional groups with populations of 94 or more at any given sub-state geography and American Indian/Alaska Native/Native Hawaiian (AIANNH) area will have data published for that geography in the 2020 Census (in contrast to the minimum population count of 100 in the 2010 Census). This means that for the 2020 Census, we should see tabulation of approximately 370 detailed racial and ethnic groups and 1,200 American Indian and Alaska Native tribes and villages. Clearly, this is more extensive than what was produced in 2010.

While there will be more detailed groups included in the data set, due to privacy concerns, the Detailed DHC-A and Detailed DHC-B will actually have fewer tables and levels of geography than in 2010. This is a significant setback. Higher levels of geography are not particularly useful when trying to identify and target areas and communities of concern – they are simply too large in many cases to provide the granularity needed. Thus, it would be ideal to produce data at the block group geography, even if it would only be for a subset of detailed groups.

---

[3] Id.

[4] https://theconversation.com/asians-could-opt-out-of-naming-a-country-of-origin-on-the-2020-census-a-policymakers-nightmare-92714.

[5]https://www.pfizer.com/news/articles/health_disparities_among_asian_americans_and_pacific_islanders#:~:text=In %20fact%2C%2019.4%25%20of%20Asian,all%20Asians%20or%20U.S.%20residents.

[6] Id.

[7] https://www.pewresearch.org/fact-tank/2021/04/29/key-facts-about-asian-origin-groups-in-the-u-s/.

It will be important for the Bureau to educate data users extensively about what to expect from these data sets, how to best use them, how not to use them, and how to best present these data. The Census Bureau must do more than releasing technical documentation with the final release of the Detailed DHC-A to achieve these goals. While technical documentation is an important standard to maintain, community-focused and data-user-focused educational materials are also needed to ensure the data are used properly and effectively by all. We must democratize access to data by making it accessible to audiences beyond technical and scientific experts .

One example of how the Bureau could better educate data users would be to address the fact that when aggregating data from the D-DHC tables, the data will generally become more variable the more they are aggregated. This will differ from other census data products and will not be intuitive to data users. The Bureau needs to do more than simply state that data users should use the published statistic they are interested in (when available) rather than aggregating data themselves. Moreover, the Bureau must go beyond warning users to exercise caution when aggregating data for custom geographies—a common practice among data users. The Bureau must provide guidance on what can and cannot be done with the data, what other methods can be utilized achieve the data of interest, and the best and most appropriate way to present the data.

Additionally, certain groups that meet specific thresholds will be eligible for age by sex data. The amount of detailed age by sex data will be based on the size of the group. At the same time, some groups will be too small to receive a sex by age table in 2020 and will only receive total population data at the national and state level. It is laudable that the Census Bureau is trying to provide as much data as possible under their adaptive design. But it will be important that the design is well explained so that people understand why they are seeing what they're seeing, including that which is hidden or excluded. That is, it will be important for people to understand why some detailed groups have age by sex data and others do not. Similarly, the Bureau needs to explain why some detailed groups have more detailed age and sex data than others. Without proper guidance and education, some may infer ill intent for why some groups have access to certain data while others do not.

Similarly, the Bureau must educate data users about the fact that the Detailed DHC-A will not be consistent with other 2020 Census data products. Providing community-focused educational materials that are easily understood will be key to ensure data users know which data products to use, for what purposes and how best to use them. This will help to ensure that census data are used appropriately and to minimize inappropriate uses of census data that could call into question the legitimacy of data collected.

Classifications of Racial and Ethnic Detailed Groups

Advancing Justice | AAJC values that the Bureau is looking to report on data for regional groups for the first time in 2020. Regional groups include groups such as Caribbean, East Asian, American Indian, Polynesian, and so forth. Because these data points were not published in 2010, this could be particularly helpful especially if additional data points are provided that would otherwise not meet the population threshold for individual detailed groups..

However, the current Race and Ethnicity Codes in Appendix F of PL94-171 Redistricting Summary File Technical Documentation and the anticipated coding for the Detailed DHC-A include an incorrect regional classification for the Hmong American community. As flagged in the sign-on letter of 57 Hmong

American; Southeast Asian American; and Asian American, Native Hawai'ian, and Pacific Islander organizations sent on March 2, 2023, Hmong is incorrectly included in the East Asian regional group. As the letter stated, "Hmong Americans' origins are as refugees from Southeast Asia [and that] Southeast Asian American (SEAA) is not only a geographic identity, but also a political identity that comes from the shared experiences of people who came to this country as refugees from Cambodia, Laos, and Vietnam." The letter further finds, "The U.S. Census Bureau's own data describes this history. According to the American Community Survey, nearly all foreign-born Hmong Americans migrated from Southeast Asia, which the Bureau defines as including the following countries: Myanmar, Cambodia, Indonesia, Laos, Malaysia, Philippines, Singapore, Thailand, and Vietnam. The most recent American Community Survey estimates show that 95.6% of Hmong Americans reported Southeast Asia as their region of origin, while only 1.6% reported East Asia." In addition to the community's understanding of Hmong as Southeast Asian Americans, this is also well-established in the academic literature.[8]

The Census Bureau must immediately rectify this error and reclassify Hmong in the Southeast Asian regional group for its data products past and future.

Need to Actively Engage Impacted Communities

The Census Bureau must do a better job of actively engaging impacted communities to ensure that their policies and decisions align with how communities self-identify. The Bureau must provide communities with the tools they need to effectively utilize census data. For example, had the Census Bureau consulted any one of the 57 groups from the sign-on letter, it would have been quite evident that the community identifies as Southeast Asian, not as East Asian. These groups include national advocacy organizations that work on behalf of, and are led by, members of the community. It will be important for the Bureau to engage with a diverse set of interested stakeholders, including academics and groups that serve the impacted communities.

The Bureau should also make sure to engage a diverse range of stakeholders as it develops community-focused educational materials that explain the Detailed-DHC datasets in easily accessible language, with extensive examples of what to do and what not to do with these data. Engaging with community groups will provide the Census Bureau an opportunity to utilize communities as informal focus groups, allowing the Bureau to fine-tune their materials to ensure they are as accessible as possible.

Conclusion

[8] For a selection, see, Wayne Carroll, "Economic Progress of Hmong Americans: The First Twenty-Five Years," Hmong Studies Journal 23 (January 2021): 1–49, accessed March 2, 2023,https://search-ebscohost-com.proxygw.wrlc.org/login.aspx?direct=true&db=a9h&AN=154703458&site=ehost-live; --- and David Schaffer, "Employment and Wages of Hmong and Other Southeast Asian Refugees in the United States," Journal of Immigrant & Refugee Studies 19, no. 4 (2021): 526–539; Bic Ngo and Stacey J. Lee, "Complicating the Image of Model Minority Success: A Review of Southeast Asian American Education," Review of Educational Research 77, no. 4 (2007): 415–453; Arthur Sakamoto, John Iceland, and Thomas Siskar, "The Socioeconomic Attainments of Second-Generation Southeast Asian Americans in the 21st Century: Evidence from the American Community Survey, 2012–2016," Population Research and Policy Review 41, no. 1 (2022): 59–88.

We appreciate the opportunity to provide comments on the 2020 Census Detailed DHC-A and the Proof of Concept. Please feel free to contact me at tminnis@advancingjustice-aajc.org or (202) 815-4412 if you have any further questions.

Sincerely,

Terry Ao Minnis
Senior Director of Census and Voting Programs
Asian Americans Advancing Justice | AAJC

Post Deadline Detailed DHC-A Proof of Concept Feedback

14. Anonymous

Mon 3/6/2023 11:06 AM

Good morning,

Thank you for sharing the proof of concept. I received your email yesterday; although my feedback seems missing the deadline, I am sending it anyway.
My review has been very superficial, but the main take aways seem that the comparability between geographies for any D or H count is not doable using the Detailed DHC-A,

1. the counts for any geography will not be comparable to any other count for the same geographic level, (Asian CT#1 vs. Asian CT#2)
2. the sum of nested geographic level cannot be aggregated to an upper geographic level. (sum of CT Asian counts to County Asian counts)

I this can help.

Please remove my identifying information when publish my feedback.

Cordially,

15. Melissa M. Smith, Centers for Disease Control and Prevention

Fri 3/10/2023 9:49 AM

Greetings Census Bureau,

Thank you for making the proof of concept materials available for review and apologies for the delay in providing these comments. I do have a note and a question.

I noticed that block and block group data was not included in the proof of concept webinar materials. We use the block level of census data routinely, so we would not be able to evaluate this product for that portion of our work. Do you know when these lower-levels of geography might be available for review?

I do have a question, regarding the guidance for data users:

- *Values from lower-level geographies may not sum to values from higher-level geographies.*
- *Data users should use the published counts for the geography of interest rather than creating custom aggregations.*

I understand that the values from the lower-level geographies may not sum to values from higher-level geographies. I am curious about the statement on custom aggregations. We routinely use custom geographic areas of study that do not conform to any other jurisdictions in our work. Specifically, we aggregate blocks together to estimate populations for an area of study. Would you have additional guidance as to how we should handle communicating uncertainties in the population estimates in this work for our custom areas of study? Are there additional calculations or margins of error or something that we would have to use? We would appreciate any additional guidance or recommendations you would have for us as we continue this work.

Thank you in advance for your assistance.

Best regards,

Melissa M. Smith, MA
GIS Specialist
Geospatial Support, Training, Analysis, and Resources (GeoSTAR) Coordinator
Geospatial Research, Analysis, and Services Program (GRASP)
Office of Innovation and Analytics, ATSDR, Centers for Disease Control and Prevention